AD-A104 733 NAVAL HEALTH RESEARCH CENTER SAN DIEGO CA F/G 5/10
AN EPIDEMIOLOGY PRIMER: BRIDGING THE GAP BETWEEN EPIDEMIOLOGY A--ETC(U)
JUL 81 L A PALINKAS, A HOIBERG
UNCLASSIFIED NAVHLTHRSCHC-81-23 NL

END
DATE FILMED
DTIC

LEVEL II

(4) AN ETHOLOGY PRIMER: BRIDGING THE GAP BETWEEN PHYSIOLOGY AND PSYCHOLOGY,

(6) [illegible]

(11) Jul 81

(12) 17

(14) [illegible]

An Epidemiology Primer: Bridging the Gap between Epidemiology and Psychology

Lawrence A. Palinkas and Anne Hoiberg*

Naval Health Research Center

P.O. Box 85122

San Diego, California 92138

Report Number 81-23, supported by Naval Medical Research and Development Command, Department of the Navy, under Research Work Unit MF58524.005-7039.

The views presented in this paper are those of the authors.

No endorsement by the Department of the Navy has been given or should be inferred.

*Environmental Medicine Department

SUMMARY

Problem

As psychologists become more involved in biomedical research issues, they are increasingly confronted by an array of methods and terms which are relatively new to them, methods and terms which lie within the province of such disciplines as epidemiology and biostatistics. The exposure to unfamiliar terms and methods can result in uncertainty and limit the effectiveness of psychologists in biomedical research.

Objective

The objective of this study was to examine the methods commonly employed in epidemiologic and biomedical research and to relate them to the methods traditionally used in the field of psychology. The intent of this report is to describe some of these methods and explain them in a clear and concise fashion.

Approach

The study was conducted in two discrete phases. In the first phase, research reports and journal articles were surveyed to ascertain the range of methods used by epidemiologists and medical researchers. These methods were placed into two categories: The first category included those statistical methods and concepts commonly used by psychologists; the second category included methods and concepts which appeared to be unique to epidemiology and biostatistics.

In the second phase, several epidemiologic and biostatistics texts were consulted and used in an attempt to explain the principles behind the methods and concepts placed in the second category. Examples which displayed the potential applications of these methods also were selected.

Results

The survey revealed many similarities in statistical methods which are quite familiar to psychologists. Hypotheses are formulated and tested in much the same manner and chi-square, regression, correlation, and analyses of variance are commonly employed in studies of morbidity and mortality.

It also was found that epidemiologic studies employ rates and measures which, although seldom seen in psychology, are based on statistical concepts and principles underlying the methods developed by psychologists. Rates such as the standardized mortality ratio and incidence and prevalence rates are measures of probability. Measures of association such as the relative risk and phi coefficient are grounded in the comparison between observed and expected frequencies on which the chi-square test employed by psychologists is based.

Conclusion

It is concluded that, despite the differences in terminology and frequent use of rates which are not found in psychology, the gap between biostatistics and psychological statistics is neither large nor complex. Because the methods of epidemiology, biostatistics, and psychology are based on common statistical principles, relatively few shifts in statistical thinking are required, other than an understanding of the terminology employed, for psychologists to attain a basic comprehension of epidemiologic findings.

An Epidemiology Primer: Bridging the Gap

between Epidemiology and Psychology

Psychologists in recent years have begun to take an active role in health-related research: to pose research questions, raise methodological issues, and furnish answers to biomedical problems. In this new role they are likely to experience some degree of confusion and frustration. Familiar ground may be obscured by new labels or subtle changes in landscape. Areas formerly traversed with confidence may now cause uncertainty or trepidation. It may be felt that biostatistics represents an entirely new set of concepts and methods that requires extensive "re-tooling" for understanding and application. How readily can psychologists familiarize themselves with this "new ground"? What shifts in statistical thinking are necessary to attain a basic comprehension of epidemiologic findings? The objective of this paper is to answer these questions by examining statistical techniques commonly used in epidemiologic studies that are generally unfamiliar to psychologists and, thus, to some extent bridge the communication gap between these disciplines.

To begin, it must be recognized that all applications, whether in the biological and health sciences or psychology, derive from the same general theory of statistics. Therefore, biostatisticians and psychologists are taught the same basic statistical concepts and methods. The existence of a common background can be shown by a comparison of textbooks in these fields which reveals very similar tables of contents including chapters on descriptive statistics, probability, the binomial distribution, the normal distribution, estimation, hypothesis testing, chi-square tests, correlation and linear regression, analysis of variance, and nonparametric methods. Epidemiologists and biostatisticians also are taught the special tools of demography and vital statistics (morbidity and mortality rates and ratios) while psychologists typically learn factor analysis and psychological scaling techniques. For the psychologist, the area of vital statistics is undoubtedly very puzzling because of the unfamiliar terminology and the unique epidemiologic perspective. Therefore, our discussion will begin with a brief description of the typical approaches taken in epidemiologic studies, specifically the nature of cross-sectional, retrospective, and prospective studies and the case-control and cohort methods. From this perspective, we will consider rates and ratios commonly employed, methods used to adjust rates so that diverse populations can be compared, and measures of association among variables that affect morbidity and mortality.

EPIDEMIOLOGIC APPROACHES TO THE STUDY OF ILLNESS

MacMahon, Pugh, and Ipsen (1960) divide epidemiologic investigations into four separate categories:

(1) Descriptive Epidemiology. Descriptive epidemiology is concerned with distributions of disease and comparisons of different populations or different segments of the same population on morbidity and mortality indices.

(2) Formulation of Hypotheses. Tentative explanations of observed disease distributions are attempted in terms of possible causal associations of a direct nature.

(3) Analytic Epidemiology. This branch of epidemiology consists of observational studies designed specifically to examine and test hypotheses developed from descriptive studies.

(4) Experimental Epidemiology. Experimental studies are conducted on human populations to confirm in a rigorous manner hypotheses that stand the test of observational analytic studies.

It seems apparent that categories (2) and (3) refer to methods of statistical inference and could readily be combined. We will treat both of these categories under the heading of analytic epidemiology where the epidemiologist or biostatistician attempts to derive inductions, generalizations, or conclusions about questions he has posed by following accepted rules of evidence. (The psychologist, of course, follows the same body of rules for arriving at conclusions also extending beyond

the immediate data.) Experimental studies in epidemiology, that is, random assignment of individuals to exposure situations or treatments, are rarely possible because of practical and ethical constraints and will not be considered further here. This paper will be concerned only with descriptive and analytic epidemiology.

DESCRIPTIVE EPIDEMIOLOGY

Descriptive statistics in epidemiology involves the use of standardized indices that reflect typical or usual values, the amount of variability in sets of observations, and relationships among variables of interest. Thus, descriptive epidemiology provides methods for organizing, summarizing, and communicating study results. Descriptive epidemiology is not concerned with the causal implications or conclusions that might be drawn from sets of data; such inferences are the province of analytic epidemiology.

The most commonly used descriptive variables in epidemiology pertain to time, person, and place. According to MacMahon et al. (1960), "Statements of the frequency of a given trait or disease manifestation in various populations are essential to descriptive epidemiology. Such statements permit comparisons between populations and between subgroups of a population with respect to the manifestation in question" (p. 51). To control for differences in population or sample size, frequencies must be expressed in the form of rates.

The calculation of rates is a simple procedure performed frequently in both epidemiological and psychological research. It is a statement of probability expressed in a quantity or degree of the phenomenon measured per unit of population per unit of time. What is known in epidemiology as a specific rate is known in probability theory as a conditional probability.

In epidemiology, this statement of probability involves three different items of information: (1) the number of persons affected by a particular illness, expressed as the numerator, (2) the population within which the affected persons are observed, expressed as the denominator, and (3) a specification of the time interval.

ANALYTIC EPIDEMIOLOGY

On the basis of a comparison of these rates in different populations or subgroups of a population, tentative hypotheses are formulated which posit causal connections between the observed distributions of disease and one or more variables or characteristics of the population. These hypotheses are then tested by specially designed observational studies.

Analytic studies are typically conducted to determine whether or not an association is present between a certain characteristic or combination of characteristics and a disease in a group of afflicted individuals. In these studies, comparisons are made between a group of persons who have the disease and a group that does not. The methods employed in these studies "depend upon observing (hence the term 'observational' studies) and quantifying whatever is being studied" (Ibrahim & Spitzer, 1979, p. 139). An example of a study in psychology of this type would be the comparison of IQs among students of different ethnic groups to determine if an association exists between race and intelligence.

Analytic studies usually fall within one of two broad categories: case-control and cohort. Both of these are referred to in the research literature by various names, thereby creating some confusion over the nature of their differences (Feinstein, 1979; Ibrahim & Spitzer, 1979; Lilienfeld, 1976; MacMahon et al., 1960). In the case-control method, affected (cases) and nonaffected (controls) groups are compared to determine whether a particular characteristic occurs with greater frequency among those affected by a certain disease or illness. There are two major forms of case-control study, retrospective and cross-sectional. In the retrospective study, the objective is to establish if the characteristic was present in the past. The investigator looks backward in time for exposure. In the cross-sectional study, the characteristic being compared is present in both cases and controls at the time of the investigation. In both types of study, analysis proceeds from effect to cause.

The second category of analytic study employs the cohort method. With this design, a particular population is examined to determine if the characteristic that may be related to the disease being investigated is present in significant quantities. The researcher looks forward in time for exposure and analysis proceeds from cause to effect. In this approach, also refer-

red to as a prospective study, the population may be followed for several years to certify which members develop or die from the disease.

There are numerous advantages and disadvantages with either research approach. Prospective studies enable the researcher to obtain direct estimates of the risk associated with a suspected causal factor and to reduce the probability of spurious relationships resulting from bias in data collecting procedures, but they also generally are laborious, time-consuming, and expensive (MacMahon, et al., 1960, p. 47). The retrospective study, on the other hand, is relatively quick and inexpensive, easily repeatable, and can economically examine a large number of cases. However, it also is subject to selection, information, and confounding biases (Feinstein, 1979; Ibrahim & Spitzer, 1979).

Whatever their relative merits and disadvantages, the objective of both types of study is to determine whether or not a relationship exists between a particular trait or set of traits and a specific illness or disease. The comparison, in its simplist, dichotomous form, is usually represented in a 2x2 table, as shown in Table 1. If a higher proportion of individuals with the characteristic is found among the cases than the controls, an association between the disease and the characteristic is indicated.

Table 1

Framework for the Study of Disease

	Number of Individuals		
Characteristic	With Disease	Without Disease	Total
With	a	b	$a+b = N_1$
Without	c	d	$c+d = N_2$
Total	a+c	b+d	$a+b+c+d = N$
	M_1	M_2	N

STATISTICAL METHODS FOR DESCRIPTIVE EPIDEMIOLOGY

Calculation of Rates

The description of a particular illness may utilize one or both of two types of rates, mortality and morbidity. To compute a mortality or death rate, the following specific information is needed: (1) In the numerator is included the number of deaths in the exposed or affected population during a certain time period. (2) In the denominator is the total population group exposed to the risk of death. (3) A time factor, usually a 1-year interval, is specified. The annual death rate can be calculated with this information.

$$\text{Annual death rate (ADR) from all causes (per 1,000 population)} = \frac{\text{Total number of deaths during a specified period of a year} \times 1{,}000}{\text{Number of persons in the population at mid-year}}$$

Thus, if 1,200 deaths occurred in a population of 1,000,000 in 1980, the annual death rate would be:

$$\text{Annual death rate (ADR) in 1980 (per 1,000 population:} = \frac{1{,}200 \text{ deaths in 1980} \times 1{,}000}{1{,}000{,}000 \text{ persons present as of July 1980}} = 1.2 \text{ per 1,000 population.}$$

The units of time and population may be selected by the investigator to suit his own purposes, but they must be specified. Death rates also can be made specific for a variety of characteristics, such as age, cause of death, marital status, race, and occupation.

$$\text{ADR from all causes for persons 18-24 (per 1,000 population)} = \frac{\text{Number of deaths of persons 18-24 during a period of 1 year}}{\text{Number of 18-24 year olds in the population at mid-year}} \times 1{,}000$$

$$\text{ADR from lung cancer (per 1,000 population)} = \frac{\text{Number of deaths from lung cancer per year}}{\text{Number of persons in the population at mid-year}} \times 1{,}000$$

Another type of rate frequently used is the "case fatality rate":

$$\text{Case fatality rate (\%)} = \frac{\text{Number of individuals dying during a specified period of time after onset or diagnosis of disease}}{\text{Number of individuals with the specified disease}} \times 100$$

This rate represents the risk of dying during a definite period of time for those individuals who have the particular disease. As with the death rate, the period of time during which the deaths occurred should be indicated. Case fatality rates also can be made specific for age, sex, severity of disease, and any other factors of clinical and epidemiological importance.

The third mortality rate used in epidemiologic research is the proportion of total deaths due to a specific cause:

$$\text{Proportionate mortality rate from cardiovascular diseases in the U.S. Navy in 1970} = \frac{\text{Number of deaths from cardiovascular diseases in the U.S. Navy in 1970}}{\text{Total deaths in the U.S. Navy in 1970}} \times 100$$

However, since this rate depends on two variables, it is of limited value in making comparisons between different populations or time periods. It also fails to directly measure the risk or probability of a person in a population dying from a specific disease as does a cause specific mortality rate.

One of the most frequently used morbidity rates is the incidence rate which is defined as the number of new cases of a disease that occur during a specified period within a specified unit of population.

$$\text{Incidence rate per 1,000} = \frac{\text{Number of new cases of a disease occurring in a population during a specified period of time}}{\text{Number of persons exposed to risk of developing the disease during that period of time}} \times 1{,}000$$

Another morbidity rate is the prevalence rate, which measures the number of cases that are present at, or during, a specified period of time. The prevalence rate equals the incidence rates times the average duration of the disease. For example, if the average duration of hypertension is three years and its incidence rate is 15 per 1,000, the prevalence rate would be 45 per 1,000.

$$\text{Prevalence rate per 1,000} = \frac{\text{Number of cases of disease present in the population at a specified time}}{\text{Number of persons in the population at that specified time}} \times 1{,}000$$

The two types of prevalence rates which are used by investigators are point prevalence and period prevalence. Point prevalence refers to the number of cases present at a specified moment in time; period prevalence refers to the number of cases that occur during a specified period of time, for example, a year. Period prevalence consists of the point prevalence at the beginning of a specified period of time plus all new cases that occur during that period.

All forms of morbidity rates, including incidence and prevalence rates can be made specific for age, sex, and/or any other personal characteristics. They also can be standardized in the same manner as mortality rates.

AGE ADJUSTMENT FOR MORTALITY RATES

The population characteristic that has the greatest influence on mortality rate is the age of the members. Since differences in the age composition of a population will influence the total mortality rates, it is preferable to use age specific mortality rates in comparing the mortality experiences in two different geographical areas, population groups, or time periods. To control for differences in the age distribution of a population, two different summary statistics may be employed: the direct method of age adjustment and the standardized mortality ratio. Both rely upon a comparison of expected rates of a standard or control group with the observed rates of the population under study.

Direct Method of Age Adjustment. The basic procedure for this method is to apply the age-specific mortality rates for the two groups that are being compared to the number in the same age groups of the standard population. For most studies conducted in the United States, the standard population is the population of the U.S. as determined in the 1940 census. This procedure gives the number of deaths that can be expected in the standard population if these age-specific rates from the observed populations had prevailed in the standard population. An example of the use of this method to adjust the calculation of a mortality rate to control for age is found in Table 2.

Table 2

Calculation of the Age-Adjusted Mortality Rates from All Causes by the Direct Method: United States, 1950 and 1960[a]

Age Group (Years)	Mortality from All Causes per 100,000 Population		Standard Population: Total U.S. Enumerated Population per 1,000,000	Expected Number of Deaths that Would Occur in Standard Population at Rates in	
	1950	1960	1940	1950	1960
	(1)	(2)	(3)	(1)x(3)	(2)x(3)
< 1	3,299.2	2,696.4	15,343	506.2	413.7
1-4	139.4	109.1	64,718	90.2	70.6
5-14	60.1	46.6	170,355	102.4	79.4
15-24	128.1	106.3	181,677	232.7	193.1
25-34	178.7	146.4	162,066	289.6	237.6
35-44	358.7	299.4	139,237	499.4	416.9
45-54	853.9	756.0	117,811	1,006.0	890.7
55-64	1,901.0	1,735.1	80,294	1,526.4	1,393.2
65-74	4,104.3	3,822.1	48,426	1,987.5	1,850.9
75-84	9,331.1	8,745.2	17,303	1,614.6	1,513.2
85+	20,196.9	19,857.5	2,770	559.5	550.4
Total death rate all ages	963.8	954.7	--	--	--
Total population	--	--	1,000,000	--	--
Total expected number of deaths	--	--	--	8,414.5	7,609.7
Age-adjusted death rate per 100,000	--	--	--	841.45	760.97

[a]Source: Klebba, Mauer, and Glass (1973)

In Table 2, the age-adjusted rate in 1960 is much lower than in 1950, in contrast to the total death rates where the 1960 death rate is only slightly lower. The difference between the changes in the total and age-adjusted death rates results from the fact that the 1960 population has a larger proportion of people in the older age groups than the 1950 population. The total death rate is affected by both the age-specific death rates and the age distribution of the population. The age adjustment procedure is used to remove the influence of the age distribution of the population by use of a standard population.

Standardized Mortality Ratio (SMR). A second method of age adjustment is a statistic widely used in studies of occupational mortality. It is defined as the number of deaths, either total or cause-specific, in a given occupational group expressed as a percentage of the number of deaths that would have been expected in that occupational group if the age- and sex-specific rates in the general population were applicable. The statistic is calculated by using the formula:

$$\text{Standard Mortality Ratio (SMR)} = \frac{\text{Observed number of deaths per year}}{\text{Expected number of deaths per year}} \times 100 = \%$$

The expected number of deaths per year of a particular sample is calculated by using the equation $N = \Sigma ab$ where

a = the number of sample members belonging to a particular age group

b = the standard death rate in the general population for that same age group.

An example of this procedure is found in Lilienfeld (1976). Between 1949 and 1953, there were 7,320 deaths among male farmers and farm managers in England, or an average of 1,464 deaths per year. In determining whether this figure indicates a normal, high, or low mortality risk for this group, a standardized mortality ratio is calculated using the standard death rates in England (see Table 3).

Table 3

Calculation of the Standardized Mortality Ratio for Occupation of Male Farmers and Farm Managers for All Causes of Death: 1951[a]

Age Group	Number of Farmers and Farm Managers (Census, 1951) (1)	Standard Death Rates per 1,000,000 (All Causes of Death) (2)	Expected Number of Deaths for Farmers and Farm Managers per 1,000,000 (3) (1) X (2)
20-24	7,989	1,383	11
25-34	37,030	1,594	59
35-44	60,838	2,868	174
45-54	68,687	8,212	564
55-64	55,565	22,953	1,275

Total expected deaths per year: 2,083

Total observed deaths per year: 1,464

$$SMR = \frac{1,464}{2,083} \times 100 = 70.3\%$$

[a]Source: Registrar General's Decennial Supplement (1958).

Table 3 gives the results of such a calculation. The SMR indicates that the mortality experience of farmers and farm managers was only 70.3% of the total population rate from all causes of death.

Although the SMR is a widely used statistic in epidemiology, it also possesses certain limitations. Wong (1977), for instance, notes that the comparison of SMRs is questionable. Gaffey (1976) cites three specific limitations to the SMR: (1) the lack of a relationship between the SMR and the life expectancy of a particular population; (2) the unequal sizes of the SMR and the relative risk, the discrepancy depending on the age of the study population, and (3) at older ages, the SMR is subject to limitations on its possible values, more or less independently of any hazard to which the study population may be exposed. Based upon these limitations, Gaffey recommends against the use of the SMR as an estimate of relative risk, believing that the SMR in general will be a biased estimate of that relative risk and its bias will be different with each age group. However, Symons and Taulbee (1981) state that the SMR can be a useful approximation of relative risk when (1) the age-specific rates in the comparison population for the cause(s) of interest are no larger than about 100 per 10,000 subjects per year, (2) the age bands are not too broad, and (3) the age-specific mortality rates for the study and comparison populations are in approximately constant ratio across the age bands.

MEASURES OF ASSOCIATION

In both retrospective and prospective studies, the object of research is to determine whether or not a correlation can be established between a specific characteristic and the disease being examined. Psychologists have long employed statistical methods, such as analyses of variance and regression analyses, to measure the strength or degree of association between the variables. In some instances, epidemiologists employ similar methods; other instances require the use of methods not found in psychology. In this paper, four specific methods commonly used in biostatistics which examine the relationship or association between two or more variables are reviewed: Poisson distributions, relative risk, measures based on chi-squares, and attributable risk.

Poisson Distributions. Hypothesis testing involves the use of a sampling distribution in which probabilities (expected frequencies) are compared with outcome (observed frequencies). A distribution of probabilities indicates the likelihood of

each of the observed frequencies if the assumptions made regarding the phenomenon under study are actually correct. There are three probability distributions commonly employed in statistical analyses: the normal distribution, binomial distribution, and Poisson distribution. Psychologists are familiar with the first two; the third, however, is more commonly found in biomedical research. Although the Poisson distribution is often used as an approximation of the binomial distributions, it may provide a useful distribution in its own right.

The Poisson distribution is used when events or entities are random and independent of each other and when the probability of an event is very small and, even if the sample size is large, only a small number of events are observed. These distributions are likely to be obtained when the observations consist of counts such as the number of cases of a particular illness over a fixed period of time.

The Poisson distribution consists of a distribution of random variables taking values of 0, 1, 2,.... If the variable X has a particular value, then the probability from the Poisson distribution that this value will occur is

$$\frac{e^{-\mu}\,\mu^{x}}{x!}$$

The quantity "e" in this formula is a constant with a value approximately equal to 2.71828. For variables that have a Poisson distribution, the mean and the variance are equal, that is to say, $\mu = \sigma$. In contrast to the binomial distribution which requires a knowledge of n and p, the Poisson distribution requires only a knowledge of the distribution mean, or μ, which can take any value greater than zero.

Remington and Schork (1970) present two general models which demonstrate the utility of the Poisson distribution. The first is characterized by a large quantity of some medium such as sea water or air in which are found a large number of small, discrete entities, such as plankton or bacteria. One of the most important traits of this model is that there is a uniform density of the small entities throughout the medium. When a small quantity of this medium is examined, the probability that this sample will contain x number of entities is the Poisson probability.

The second model producing Poisson probabilities concerns events occurring in time. Such an event would include members of a community who contract a particular illness or disease. If the events occur independently, the probability that an event will occur in a short-time interval is proportional to the length of the interval, and the time interval is short enough such that the probability of more than one event occurring in such a time interval is negligible, then the probability that x events occur in a fixed time interval is the Poisson probability.

Relative Risk. The most common measure of association in retrospective studies is relative risk which reflects the incidence of disease among a group possessing a certain characteristic relative to a group without the characteristic. The measure indicates the likelihood a member of a specified population will acquire and/or succumb to a disease if he possesses the characteristic under study. Thus, a study which determines that the relative risk of lung cancer for cigarette smokers is 3.3 is stating that the risk of contracting lung cancer is 3.3 times greater for smokers than for nonsmokers.

Relative risk is calculated from a 2x2 table in which the number of cases and controls are compared with respect to the presence or absence of a particular characteristic (see Table 1). The cross products are then multiplied and divided, producing the following equation:

$$RR = \frac{ad}{bc}$$

This equation gives an approximation of relative risk and assumes that (1) the cases and controls have been selected at random and are representative of the larger population, and (2) the frequency of the disease in a population is relatively small. If RR is equal to 1 or unity, then $\frac{ad}{bc}$ as an approximation of relative risk is exact. This equation is the one used most often in calculating relative risk.

If the frequency of disease in a population is large or the approximation of RR proves to be inadequate, i.e., in cases

where there are multiple categories of groups--different subgroups by age or occupation--under study, there is a more accurate measure developed by Mantel and Haenszel (1959). The revised relative risk is calculated as follows:

$$RR_{mh} = \frac{\Sigma(\frac{ad}{N})}{\Sigma(\frac{bc}{N})}$$

In addition, Mantel and Haenszel have calculated summary relative risk equations for separate subcategories of exposure. The rationale for these equations is that "...over-all relative risk estimates are averages and as averages may conceal substantial variation in the magnitudes of the relative risk among subgroups. Ordinarily, the individual subcategory data should be examined, paying special attention to relative risks based on reasonably large sample sizes. This will provide protection against the potential deficiencies of any particular summary relative risk formula employed" (Mantel & Haenszel, 1959, p. 740). An example of such a summary relative risk formula is the following:

$$RR = \frac{\frac{ad}{bc}}{\frac{E(a)E(d)}{E(b)E(c)}}$$

where:

$$E(a) = \frac{N_1M_1}{N}$$

$$E(b) = \frac{N_1M_2}{N}$$

$$E(c) = \frac{N_2M_1}{N}$$

$$E(d) = \frac{N_2M_2}{N}$$

A simpler method of calculating relative risk for multiple categories, however, is to prepare a series of 2x2 tables comparing controls and cases at different levels of exposure and then to compute the relative risk for each table. An example of such a procedure is found in Lilienfeld (1976) (see Table 4).

Age adjustment procedures are also important when calculating relative risk. One such procedure is the matching case method in which a sample of N diseased individuals is drawn and the characteristics of each individual noted with respect to the control factors. Subsequently, a sample of N well individuals is drawn, with each individual matched on the control factors to one of the diseased individuals. In applying such a procedure, the 2x2 table takes on a different form from that shown in Table 4. The cell in Table 5 in the upper left-hand corner contains r number of pairs in which both cases and controls possess the characteristic of interest. The marginal totals (a,b,c,d) represent the entries in the cells of Table 1 and the total for the entire table is $\frac{1}{2}N$ pairs where N represents the total number of paired individuals. The calculation of the relative risk for this table would be:

$$RR = \frac{S}{t} \text{ (provided } t \neq 0)$$

Table 4

Relative Risk for Smokers and Nonsmokers

Example of Calculating Relative Risk for Multiple Categories

Daily Average Cigarettes Smoked	Patients		Relative Risk of Different Categories of Smokers to Nonsmokers
	Lung Cancer	Controls	
0	7	61	1.0
1-4	55	129	3.7
5-14	489	570	7.5
15-24	475	431	9.6
25-39	293	154	16.6
50+	38	12	27.6

The different degrees or levels of cigarette smoking are to be compared with the nonsmokers, and, therefore, the relative risks of lung cancer for nonsmokers is taken to be 1.0. The risks for smokers compared to nonsmokers are:

RR (1-4 cigarettes daily) $\frac{55 \times 61}{\sim \times 129}$ $\frac{3,355}{903}$ 3.7

RR (5-14 cigarettes daily) $\frac{489 \times 61}{\sim \times 570}$ $\frac{29,829}{3,990}$ = 7.5

RR (15-24 cigarettes daily) $\frac{475 \times 61}{\sim \times 431}$ $\frac{28,975}{2,917}$ 9.6

Table 5

Model of Calculation of Relative Risk for Matched Cases and Controls With and Without a Characteristic

	Controls		
Cases	With Characteristic	Without Characteristic	Total
With characteristic	r	s	a*
Without characteristic	t	u	c*
Total	b*	d*	$\frac{1}{2}$N

*a,b,c, and d are the entries in the cells in Table 1.

A test of whether or not the observed difference between ad and bc is due to sampling variation is provided by a chi-square test for 2x2 tables. Mantel/Haenszel have developed a chi-square formula specifically for use in testing the significance of a relative risk correlation.

$$\chi^2_{mh} = \frac{(|\Sigma a - \Sigma E(a)| - \frac{1}{2})^2}{\Sigma V(a)}$$

where $E(a) = \frac{N_1 M_1}{N}$

and $V(a) = \frac{N_1 M_1 N_2 M_2}{N^2(N-1)}$

If $\chi^2 > 3.84$, one may conclude that it is unlikely that the difference in risk between the group with and the group without the characteristic is a result of chance.

When testing for significance in a matched pairs example such as in Table 5, the McNemar test where:

$$\chi^2 = \frac{(|t-s|-1)^2}{t+s} \quad \text{with 1 df}$$

may be employed.

In establishing confidence limits for the test of significance, confidence limits of the logarithm (to the base e) of a corrected relative risk are computed and the logarithmic confidence levels are then reconverted to the original scale. The addition of 0.5 to the numbers a,b,c, and d corrects for a bias which can occur with small numbers of observations. Using the log-relative risk rather than the relative risk itself simplifies calculations of standard errors necessary for computing confidence intervals.

Using the Chi-Square Test. The chi-square, a statistical tool familiar to most psychologists, has two basic uses in epidemiologic research. First, the chi-square test may be used to evaluate whether or not frequencies which have been empirically obtained differ significantly from those which would be expected under a certain set of theoretical assumptions, that is, testing the null hypothesis. Second, the chi-square test may be used in determining the degree or strength of an

association. As with the standardized mortality ratio, the chi-square is based upon a comparison of observed and expected frequencies. Chi-square is obtained by taking the square of the difference between the observed and expected frequencies in each cell divided by the expected number of cases in each cell:

$$\chi^2 \quad \sum \frac{(f_o - f_e)^2}{f_e}$$

$\chi^2 = 0$ when all observed and expected frequencies are identical. If $f_o - f_e \quad 0$, then the null hypothesis is confirmed. The greater the difference between observed and expected frequencies, the greater the chance the null hypothesis is rejected.

The chi-square for a 2x2 table may be calculated using the formula

$$\chi^2 = \frac{(|O-E|-\frac{1}{2})^2}{E} = \frac{(|ad-bc| - N/2)^2 N}{(a+b)(b+c)(a+c)(d+b)}$$

In the case of a 2x2 table where any of the expected frequencies are 5 or less, a correction for continuity, known as the Yates correction, can be made by adding or subtracting 0.5 from the observed frequencies in order to reduce the magnitude of the chi-square. The usefulness of this modification, however, has been subject to debate (cf., Fleiss, 1973; Mantel and Greenhouse, 1968; Grizzle, 1967; Remington & Schork, 1970).

A common fallacy in employing the chi-square test is to use the value of chi-square itself as a measure of the degree to which a disease and a characteristic are associated with one another. Even though chi-square is excellent as a measure of the significance of an association, it does not indicate the degree of association because it is a function both of the properties of the various cells and the total number of subjects studied. The degree of association present is really only a function of the cell proportion, which explains why relative risk is used as a measure of association.

There are, however, measures based upon the chi-square test which do provide a measure of the degree of the association between an illness and a specific characteristic. One such measure is the phi coefficient. The phi coefficient or ϕ gives a numerical value, ranging from 0 to +1 for a relationship between two variables and is similar in meaning to a correlation coefficient. It is calculated by using the following formula:

$$\phi \quad \frac{(ad-bc)}{\sqrt{(a+b)(a+c)(b+c)(b+d)}} = \frac{(ad-bc)}{\sqrt{N_1 N_2 M_1 M_2}} = \frac{\chi^2}{N}$$

Another measure is V or Cramer's measure and is calculated by using the formula:

$$V^2 = \frac{\chi^2}{N \text{Min}\,(r-1, c-1)} = \frac{\phi}{\text{Min}\,(r-1, c-1)}$$

where Min (r-1,c-1) refers to either r-1 or c-1, whichever is the smaller.

Another measure is the Pearson's contingency coefficient where:

$$C = \frac{\chi^2}{\chi^2 + N}$$

Attributable Risk. A fourth measure of the association between a disease and a particular characteristic is "attributable risk." The measure was initially defined in terms of lung cancer and smoking as the maximum proportion of lung cancer attributable to cigarette smoking (Levin, 1953). It is expressed as:

$$AR = \frac{b(r-1)}{b(r-1)+1}$$

where r = relative risk of lung cancer among cigarette smokers as compared to nonsmokers, and b = proportion of the total population classified as cigarette smokers.

The effect of relative risk (r) and the proportion of those with a characteristic in the population (b) on the values of the attributable risk are shown by calculations of the attributable risk for different values of r and b in Table 6. Thus, "...when the frequency of a characteristic, such as cigarette smoking is low and the relative risk for a disease among ciga-

rette smokers is also low only a small proportion of the cases of disease can be attributed to cigarette smoking" (Lilienfeld, 1976, p. 185). The reverse is true when relative risk and the proportion of smokers is high. Attributable risk, therefore, allows one to estimate the extent to which a particular disease is due to a specific factor.

Table 6

Attributable Risks as a Proportion for Selected Values of Relative Risk and Proportion of Population with the Characteristic

b = Proportion of Population with Characteristic (percent)	r Relative Risk			
	2	4	10	12
10	.09	.23	.47	.52
30	.23	.47	.73	.77
50	.33	.60	.82	.84
70	.41	.67	.86	.89
90	.47	.73	.89	.91
95	.49	.74	.90	.92

Attributable risk is particularly useful for the study of mortality. For the study of fertility or recurrent diseases, however, the measure is limited because the relative risk involved is the ratio of two probabilities (Park, 1981). Park provides a modification of the attributable risk measure that is suitable for recurrent events.

SUMMARY

Even a brief survey of epidemiologic and biomedical research reveals the use of statistical methods quite familiar to psychologists. Hypotheses are formulated and tested in much the same manner and chi-squares, regression, correlation, and analyses of variance are commonly employed in the effort to study the relationships between morbidity, mortality, and numerous other environmental, physiological, social, cultural, and psychological variables.

There are, however, key statistical concepts widely used in epidemiology and biostatistics but seldom seen in psychology. Epidemiologic studies employ the use of rates and measures of association which indicate the degree of a relationship between a disease and one or more characteristics. Both the rates and the measures of association are based upon statistical concepts and principles underlying the methods utilized by psychologists. Rates, such as the standardized mortality ratio and the incidence rate, are measures of probability in which a group of people is compared with the larger population over a specified period of time. Measures of association such as the relative risk and phi coefficient are grounded in the comparison between observed and expected frequencies on which the chi-square test employed by psychologists is based.

Despite the differences in terminology and the frequent use of rates in epidemiologic and biomedical research which are not found in psychology, this brief review indicates that the gap between biostatistics and psychological statistics is neither large nor complex. Psychologists should be able to readily familiarize themselves with this new ground as the concepts which underlie the methods of epidemiology and biostatistics are also found in the statistical methods of psychology. Other than an understanding of the terminology employed, relatively few shifts in statistical thinking are necessary to attain a basic comprehension of epidemiologic findings.

REFERENCES

Feinstein, A. R. Methodological problems and standards in case-control research. Journal of Chronic Disease, 1979, 32, 35-41.

Fleiss, J. L. Statistical Methods for Rates and Proportions. New York: John Wiley & Sons, 1973.

Gaffey, W. R. A critique of the standardized mortality ratio. Journal of Occupational Medicine, 1976, 18, 157-160.

Grizzle, J. E. Continuity correction in the X^2 test for 2 x 2 tables. American Statistician, 1967, 21, 28-32.

Ibrahim, M. A., & Spitzer, W. O. The case control study: The problem and the prospect. Journal of Chronic Diseases, 1979, 32, 139-144.

Klebba, A. J., Mauer, J. D., & Glass, E. J. Mortality trends: Age, color, and sex, United States 1950-69. National Center for Health Statistics, Vital and Health Statistics Series 20, No. 15. Washington, D.C.: Public Health Service, U.S. Government Printing Office, 1973.

Levin, M. L. The occurrence of lung cancer in man. Acta Unio International Contra Cancrum, 1953, 9, 531-541.

Lilienfeld, A. M. Foundations of Epidemiology. New York: Oxford University Press, 1976.

MacMahon, B., Pugh, T. F., & Ipsen, J. Epidemiologic Methods. Boston: Little, Brown & Co., 1960.

Mantel, N., & Greenhouse, S. W. What is the continuity correction? American Statistician, 1968, 22, 27-30.

Mantel, N., & Haenszel, W. Statistical aspects of the analysis of data from retrospective studies of disease. Journal of the National Cancer Institute, 1959, 22, 719-746.

Park, C. B. Attributable risk for recurrent events: An extension of Levin's measure. American Journal of Epidemiology, 1981, 113, 491-493.

Registrar General's Decennial Supplement. 1951 Occupational Mortality, Part II, Vol. 2. London: Her Majesty's Stationery Office, 1958.

Remington, R. D., & Schork, M. A. Statistics with Applications to the Biological and Health Sciences. Englewood Cliffs, N.J.: Prentice Hall, 1970.

Symons, M. J., & Taulbee, J. D. Practical considerations for approximating relative risk by the standardized mortality ratio. Journal of Occupational Medicine, 1981, 23, 413-416.

Wong, O. Further criticisms on epidemiological methodology in occupational studies. Journal of Occupational Medicine, 1977, 19, 220-222.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER 81-23	2. GOVT ACCESSION NO. AD-A104733	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) An Epidemiology Primer: Bridging the Gap between Epidemiology and Psychology		5. TYPE OF REPORT & PERIOD COVERED Interim 6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Lawrence A. Palinkas and Anne Hoiberg		8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS Naval Health Research Center P.O. Box 85122 San Diego, California 92138		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS MF58524.005-7039
11. CONTROLLING OFFICE NAME AND ADDRESS Naval Medical Research and Development Command National Naval Medical Center Bethesda, Maryland 20014		12. REPORT DATE July 1981 13. NUMBER OF PAGES 14
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office) Bureau of Medicine and Surgery Department of the Navy Washington, D.C. 20372		15. SECURITY CLASS. (of this report) UNCLASSIFIED 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE
16. DISTRIBUTION STATEMENT (of this Report) Approved for public release; distribution unlimited.		
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)		
18. SUPPLEMENTARY NOTES		
19. KEY WORDS (Continue on reverse side if necessary and identify by block number) Biostatistics Epidemiologic methods Relative risk Standardized mortality ratio Attributable risk Rates and ratios		
20. ABSTRACT (Continue on reverse side if necessary and identify by block number) As psychologists increasingly become involved in biomedical research, it is crucial that they understand the methods of biostatistics and epidemiology. Oftentimes, a difference in terminology or methodology unique to epidemiologic research ha ers the psychologist's efforts in addressing biomedical issues. This paper presents a precis of biostatistical and epidemiologic methods that might be unfamiliar to most research psychologists and attempts to explain the usefulness of mortality and morbidity rates and ratios, age-		

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-LF-014-6601

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

adjusted rates and measures of association such as relative risk, coefficients based on chi-squares, and attributable risk. From this explanation of epidemiologic techniques, it is concluded that the transition from psychological statistics to biostatistics and epidemiology requires little "re-tooling" for understanding and application.

END

DATE

FILMED

11-81

DTIC